Platform Conference

AMD's x86-64 Technology: Extending the x86 ISA to 64-bits

Wayne Meretsky
AMD Fellow
Computation Products Group

Agenda

- AMD’s Eighth-Generation Overview
- Why is 64-bit computing important now?
- The x86-64 architecture
- Comparison to Existing 64-bit Architectures

Platform Conference

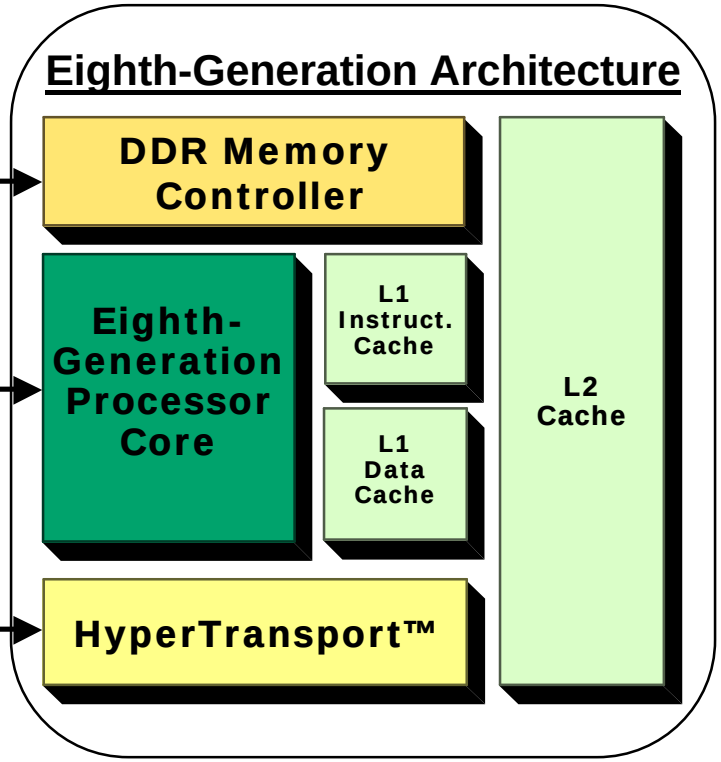
AMD Opteron™ and Eighth-Generation AMD Athlon™

Processor Architecture Overview

Key Architecture Innovations

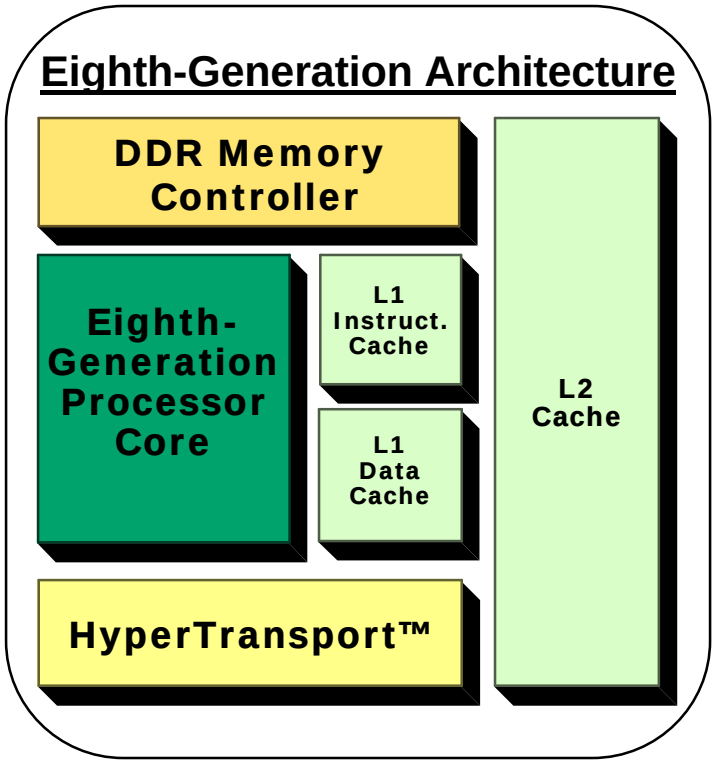
- AMD's eighth-generation architecture integrates key system elements

- **Integrated DDR memory controller**
- **Eighth-generation core**
- **HyperTransport™ technology**



Eighth-Generation Processor Core

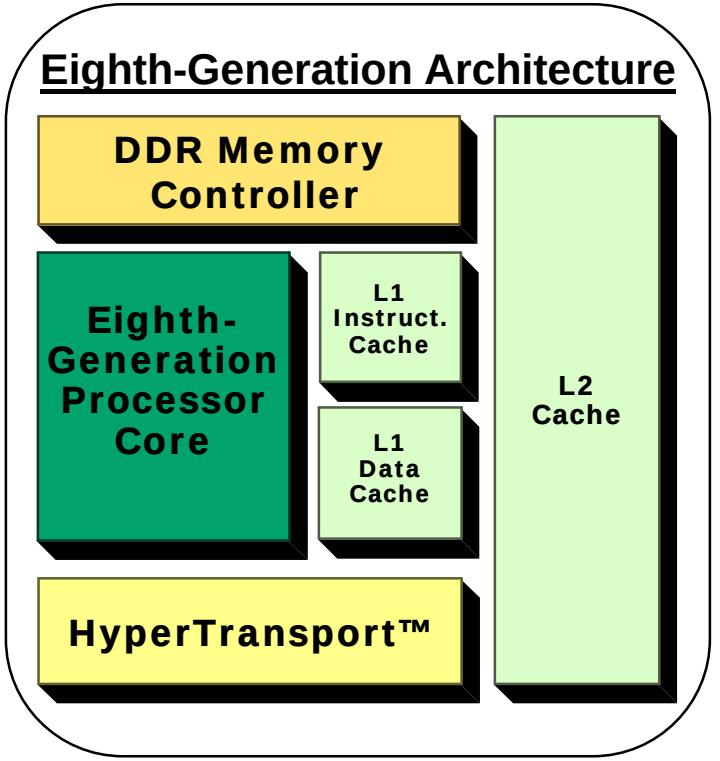
- What is AMD's eighth-generation processor core?
  - AMD's latest processor innovation, which is designed to support both 64-bit and 32-bit computing at exceptional performance levels
  - Processor core details:
    - Support for x86-64 technology
    - 12-stage superscalar pipeline
    - Big workload features
      - Enlarged 2-level TLB
      - TLB flush filter
      - Enhanced branch prediction
      - Large caches w/improved bandwidth
    - Reliability features
      - ECC protection on large arrays
      - Parity protection on major busses



- What does it do?
  - Designed to provide high performance and high throughput on both 32-bit and 64-bit applications
  - Special attention to system reliability

Integrated Memory Controller

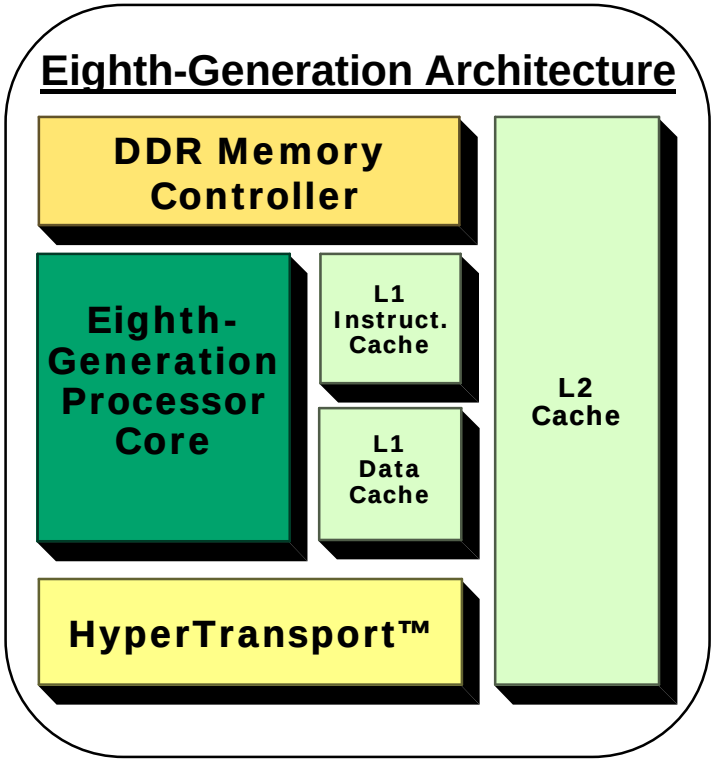
- What is an Integrated Memory Controller?
  - Integration of memory controller functionality into the processor
    - Traditionally a function of the North Bridge
  - DDR memory controller details
    - Single or dual channel DDR memory interface
    - Registered or Unbuffered DIMMs
    - Support for PC1600, PC2100, or PC2700 DDR SDRAM
    - Full ECC & Chip Kill for reliability
    - High bandwidth (up to 5.3GB/s)



- What does it do?
  - Memory controller operates at processor frequency dramatically reducing latencies
  - Memory latencies scale with processor speed
  - Memory bandwidth and capacity scales as the number of CPUs increase

HyperTransport™ Technology

- What is HyperTransport™ Technology?
  - HyperTransport technology is a high-speed, low pin-count, asynchronous, point-to-point link connecting chips
  - Highly scalable to support 1P, 2P, 4P, and 8P multiprocessing system architectures
  - Flexible
    - One, two, or three links
    - 2, 4, 8, 16, or 32-bits wide
  - High bandwidth
    - Up to 6.4 GB/s bandwidth per link
  - In production today
- What does it do?
  - Designed to increase overall performance by:
    - Removing I/O bottlenecks
    - Increasing bandwidth
    - Reducing latency
  - Enables glueless multiprocessing
  - Easy building-block approach to system design
  - Compatible with high-volume board design rules



Platform Conference

Why is 64-bit computing important now?

Interesting Trends to Note

- Gigabit DRAM devices seem likely to become mainstream in 2005
  - Single DIMM is 2 gigabytes
- Enterprise focus on integrated systems (e.g. ERP) require processor ability to directly address vast amounts of data to quickly and efficiently cull meaningful information
- Consumer access to broadband communication implies huge volumes of data flowing into the home
- Scale-out technology appears likely to overtake scale-up approaches
- The PC processor has a long history of being successfully extended to stay in front of emerging trends

Environmental Scan for 64-bit Motivators

- Eighth-generation design incorporates significant topology changes
  - HyperTransportTM connectivity
  - Integrated memory controller
- Eighth-generation designed to deliver greater performance than previous PC processors
- Trends indicate that 64-bit computing will become mainstream in all segments
- Conclusion: Eighth-generation should enable 64-bit computing

64-bit Architectural Considerations

- Required for large memory programs
  - Aggregate workloads
  - Large databases, analytics, modeling
  - Scientific and Engineering Problems
    - Designing CPUs
- But,
  - Limited Demand for Applications which require 64 bits
    - Most applications can remain 32-bit x86 instructions, if the processor continues to deliver leading edge x86 performance
- And,
  - Software is a huge investment
    - Operating systems, tools, diagnostics, applications, certifications, training, etc.
  - Instruction set is first and foremost a vehicle for compatibility

Immutable Laws of the PC Industry

- The "Immutable Laws" of the PC industry surrounding successful technology transitions are:
  - New product must cost less while providing equal functionality/performance to previous generation
  - Or …
  - New product must provide greater functionality/performance at a cost comparable to previous generation
  - And …
  - New product MUST maintain legacy compatibility
- Violators of the "Immutable Laws" perish

Platform Conference

x86-64 Technology: Extending the x86 ISA

Simple Approach

- Preserve the investment already made in x86 software
  - Full compatibility with existing 32-bit application software
- Add 64-bit capabilities to the x86 ISA, the largest software installed base
  - Provide a seamless, flexible migration path to 64-bit computing
- Leverage pricing structure and supporting software infrastructure of x86 architectures
  - Easier for ISVs
  - Decades of experience and development of x86 applications
- Innovate with open standards, compatibility, and a focus on our customers
  - Ease of implementation
  - Ease of deployment
- Leverage existing x86 dominance in the server segment by focusing on 1P to 8P solutions
- Introduce competition and provide innovation for the enterprise 32-bit and 64-bit environments
  - Provide our customers with a choice
    - Differentiated solutions provider
    - Seamless, flexible migration path to 64-bit computing
  - Enable a low-cost infrastructure for our customers by leveraging PC desktop volumes

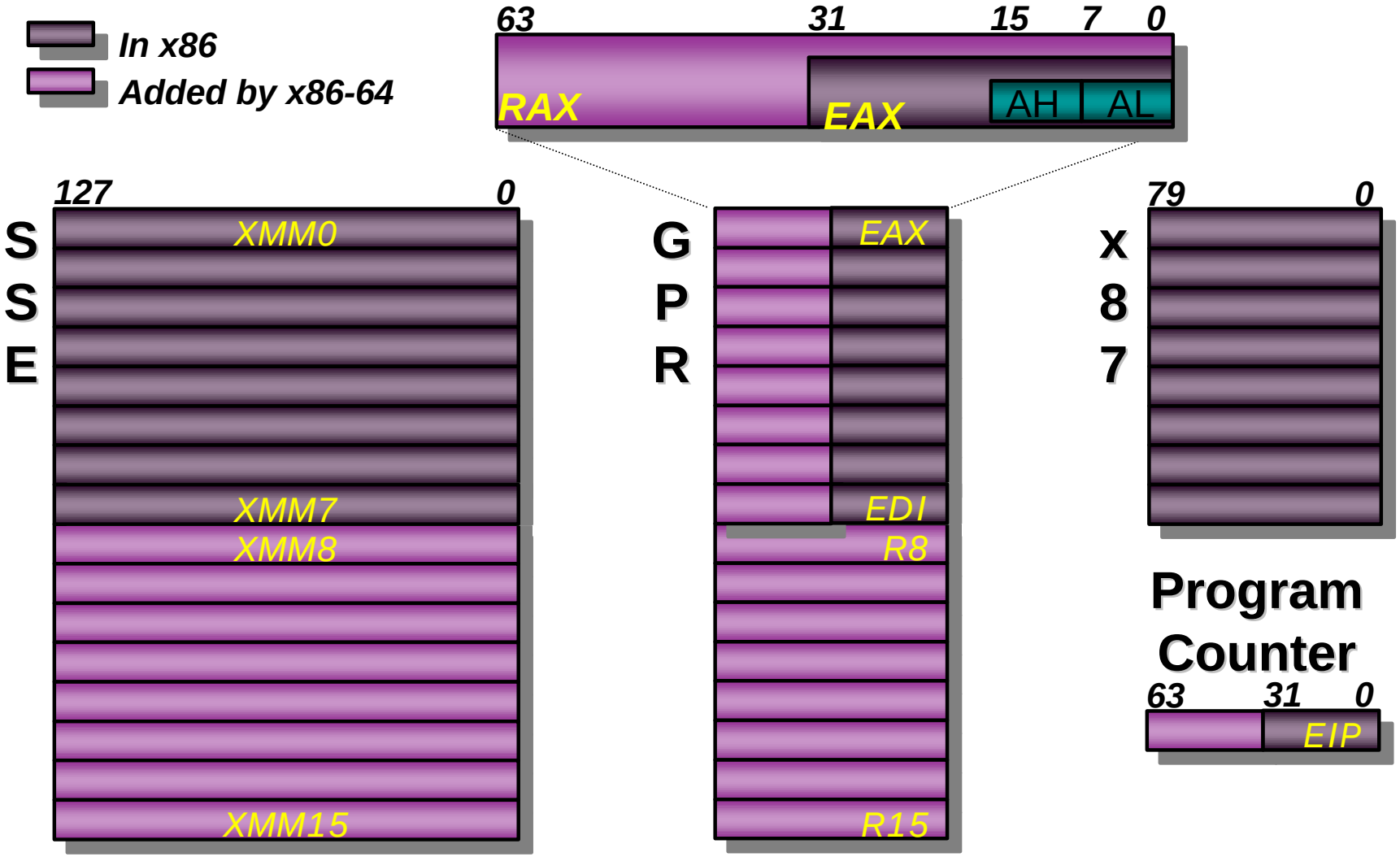
x86-64 Instruction Set Architecture

- x86-64 ISA used by placing processor in “long” mode
  - Similar to the previous extension from 16-bit 286 to 32-bit 386
  - Defaults in this mode are
    - 32-bit data
    - 64-bit addressing
  - Override prefixes allow
    - 16-bit data
    - 64-bit data
    - 32-bit addressing

- Vast majority of opcodes and features unchanged
  - Instruction sequences for 64-bit code are often identical to 32-bit sequences

- Enhancements
  - Add 8 additional general purpose registers
  - Add 8 additional floating point registers
  - Add PC-relative addressing mode
  - Add two new instructions

x86-64 Programmer's Model



Other architectural additions

- PC relative addressing mode
  - Allows Position Independent Code
    - Eliminates most load-time relocation of code
  - Tends to reduce code footprint
  - Simplifies shared code considerations
- New instructions – only two
  - MOVSXD
    - Move sign extended 32-bit to 64-bit
  - SWAPGS
    - Kernel mode instruction designed to reduce overhead in interrupt handlers
- Maintain stack alignment at 64-bits
- Reclaim some redundant instruction encodings
- Public specification available at www.x86-64.org

ABI Leverages Architectural Enhancements

- Optimized for
  - Performance
  - Fewer memory references
  - Smaller code
- Parameter and results passed via registers
  - RISC-like conventions
  - Fewer memory references required for call/exit
  - More compact prolog, epilog and call sites
- Floating point migrates to SSE registers & instructions
  - No more x87 usage
  - Register-based model allows classic optimization for float
  - Enables better super-scalar scheduling of execution units

Architectural Summary

- Instruction set is designed to offer the advantages of both CISC and RISC
  - Code density of CISC
  - Register usage and ABI models of RISC
  - Enables easy application of standard compiler optimizations
- 32-bit compatibility is excellent and easy to deliver
- Adoption is greatly simplified because of similarity to x86
  - Tool development is straightforward
  - Engineers learn new architecture very easily
  - Windows® and Linux ports have come up very quickly

Architectural Feedback*

- Code size is up about 5%
  - Mostly due to 64-bit literals
- Instruction count is down about 15%
  - Additional registers really paying off
  - Many spill/fill memory references eliminated
  - Call-Exit sequences vastly improved
- Reduced instruction count and increased IPC provides substantial performance gains
  - 8th Generation Processor Core improves IPC about 5%
  - x86-64 instruction count down about 15%
  - Net improvement of about 20% for integer code
  - All without any specific code optimizations
  - Your mileage will vary

- * Based on SPECInt2000 and pre-released x86-64 GCC compiler

Comparison to other 64-bit Architectures

The Big Four

- Four primary choices are in production today
  - Compaq's Alpha® with TRU64
  - HP's PA-RISC with HP/UX
  - IBM's Power4 with AIX
  - Sun's UltraSPARC with Solaris
- None ships in PC-processor volumes
  - Systems are very expensive
  - Software is not widely available
  - Revenue streams do not allow sufficient investment in processor development to compete with PC-processor performance
- The markets for these architectures are being cannibalized today by the PC-processor's economic proposition

What about IA-64?

- IA-64 architectural feedback is opposite that of x86-64
  - Code size is up
  - IPC is down
  - Developer learning cycles are very long
- IA-64 requires extensive compile-time optimization
  - Opportunities for parallelism are far harder to detect at compile-time than at run-time
  - Divergent from the direction suggested by both JAVA and .NET
- Lack of compatibility with x86 binaries precludes IA-64 from mainstream usage
  - Platforms, tools and software will not be available in the quality, volumes or price points associated with high-volume PC business
  - IA-64 is a late comer competing with the Big Four

Benefits of the AMD Solution

- Full 32-bit application performance
- Compelling 64-bit migration strategy
  - Gradual migration to 64-bits at customer's own pace
  - Only applications that require 64-bits need be ported
  - Linux and Microsoft® Windows® support
  - No recompiling needed for existing 32-bit installed software base
- Investment protection
  - Full compatibility with existing 32-bit software
    - Operating systems, applications, tools, diagnostics, drivers, etc.
  - Preserve investments in personnel certifications and training
- Server, Workstation, Desktop, and Mobile share same architecture
  - OS, Drivers and Applications can be the same
  - CPU vendors focus not split, ISV focus not split
  - Support, optimization, etc. all designed to be the same
- Lower Total Cost of Development and Ownership (TCD/TCO)
  - New, competitive choices for enterprise computing solutions
  - Continued innovation
  - Reduced development complexity
  - Enterprise supply chain competition
- Creates new market opportunities for our customers and partners

Cautionary statements